PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Quality improvement in a crisis: A qualitative study of experiences and lessons learned from the Irish National Ambulance Service response to the COVID-19 pandemic
AUTHORS	Heffernan, Eithne; Keegan, Dylan; Clarke, Bridget; Deasy, Conor; O'Donnell, Cathal; Crowley, Philip; Hughes, Angela; Murphy, Andrew; Masterson, Siobhan

VERSION 1 – REVIEW

REVIEWER	Yusuke Katayama Osaka University, Department of Traumatology and Acute Critical Medicine
REVIEW RETURNED	28-Sep-2021

GENERAL COMMENTS	Thank you for the opportunity of reviewing this manuscript about Irish National Ambulance Service response to the COVID-19 pandemic. I was very interesting in reviewing this manuscript, but this manuscript had many problems for publication. Major point First of all, most of the readers do not fully understand the COVID-19 pandemic in Ireland. Therefore, the authors should describe the COVID-19 pandemic in Ireland in detail. Secondly, the authors should describe how many COVID-19 response teams there are in Ireland. Without this data, it is impossible to determine if the number interviewed in this study is appropriate. Before conducting this study, did you make an estimate of the minimum number of people who the author would need to interview? Even if the number of COVID-19 response team members is small, the number of people interviewed was small, and the authors should probably have more participants in this study in order to generalize the results. Third, although it was stated that the COVID-19 response team includes a general practitioner, there was no general practitioner among those interviewed in this study. All professions should be included in the study in order to evaluate the feedback from the interviews. Finally, the authors did not clearly describe the conclusion. If the conclusion is not clearly described, the readers cannot understand what the authors would like to convey in this study. The authors should clearly describe what they would like to convey, including what is important in such a health crisis. Minor point In Page 8, the phrase “it has been reported in line with the

	Standards for Reporting Qualitative Research" should be written at the end of "design" section. In Page 8, the authors described "The full protocol has been published previously.", but the authors should cite the literature about the full protocol.

REVIEWER	Nigel Rees Welsh Ambulance Service NHS Trust, Pre Hospital Emergency Research Unit
REVIEW RETURNED	07-Oct-2021

GENERAL COMMENTS	Congratulations to the team in gathering evidence during this time and developing the body of knowledge in this area. There is a good discussion around the role of QI and improvement science methodologies. It is also argued that the COVID-19 has prompted the development of quality improvement (QI) methods and tools suited to evaluating pandemic-imposed change, such as rapid learning cycles and after-action reviews. Despite this it is difficult to find evidence of this within the papers cited and those which are cited such as Fitzsimons [1] appear to lack such scientific rigor and include opinion pieces. It is also difficult to determine the scientific quality of such approaches when the researchers highlight how in times of crisis, QI approaches must be adapted and streamlined so that improvements are implemented quickly without unnecessarily burdening strained services and staff. The paper may benefit from considering the validity of such approaches and implications of adapting or streamlining such approaches? A counter argument to such a position has been made in many articles which report poorly governed projects and low-quality methodologies employed during this time such as discussed by Lipworth et al (2020) [2]. The researchers tell how the research paradigm was pragmatism, and they report how the study was a component of a mixed methods evaluation of the NAS response to the pandemic. Whilst they provide Johnson et al (2004) [3] as a reference for this, it would help to be cited on the full study. Was the protocol prospectively published? Or are any outputs in the public domain? There is a good description on the research methods, development, and review of the interview schedule, and aim for data saturation. There is also a good description of the sample. I was however concerned with participants being recruited via a study invitation email sent by a NAS manager to former and current staff. Has the team considered their anonymity in this process and if participants were indeed open to discuss freely? There is a good description of background of researchers, which is important in their chosen methodology. They have also told how they maintained field notes and reflexive notes throughout. The use of Braun and Clark is also appropriate. This along with the peer debriefing enhanced the trustworthiness of the analysis. The researchers however mentioned how the interview schedule was reviewed and, where necessary, refined following each interview (Supplementary File 2). It would help if they could justify and describe further methodology used. The researchers told how the research was underpinned by the HSE Framework for Improving Quality and that it was used to examine the utility of this framework in a pandemic setting and inform the preparation of health services, particularly Emergency Medical Services (EMS), for future crises. Whist it seems acceptable to draw on this framework, the study in my view is significantly undermined

	by the influence of this framework. I would encourage the researchers to consider and provide a discussion on the prominent role of this framework and implications of emergence of themes versus forcing the data and confirming knowledge within this framework. For instance, many of the quotes provided do not seem to fit this framework but are coded within these themes. The Leadership for Quality theme is one such example as the following quote talks about management, performance, and wellbeing, yet there is little consideration of these issues within the discussion or how they relate to leadership? "How it's managed: That's the crux of it all…You can have amazing workers…but if it's not managed with respect and…caring, people won't…get the job done…because…it's just like a pressure cooker," (Interviewee 7). The following quote was also under the theme of Use of improvement methods. Despite this, there is little evidence in the quote of improvement methods/science/models or how they were adapted, but rather it seems to refer to leadership: Many reported that the response room manager provided strong leadership: Nevertheless, improvements were possible: "The clinical team…towards the end of the first peak...created a few standard operating procedures and we created lists…with all the important contacts [and] what we thought worked, what we thought didn't work...That was a really invaluable document…if they were training anyone," (Interviewee 8). The following quote also undermines some of argument made in the paper around improvement and rigorous QI approaches: "During the first wave, there was very little time to…look back and see how things could be improved. It was more so a case of…just fighting against the wave…certainly from our level," (Interviewee 9). N Many authors have highlighted how much of the experiences of the COVID-19 pandemic had been endured by others during previous pandemics, were somewhat predictable and within the published literature. Rather than having little time to look back, it has been argued that reflecting on/reviewing literature from previous pandemics should have been essential. There is again a mismatch between the theme The measurement for quality. There is no evidence of improvement science/QI models or their adaptations within the quote. This theme and resulting discussion may therefore have been better served by exploring the lack of rigor, validity, definition, and deployment of QI models in the EMS during the pandemic? In my view, the theme Infrastructure for quality is an example to the rest of the study themes, as they appear to suggest that it emerged from the data rather than being forced into a framework which already existed. Again however, much of the quotes may relate to something else other than infrastructure, such as communication and IT? Within the paper many inferences do not seem to be justified by this study, such as: In particular, the findings suggest that there are seven QI drivers that should be addressed when preparing for and responding to pandemics and similar emergencies. And The results underscore the importance of empathic and respectful leadership that comprises regular and clear communication and consultation with staff, which aligns with best practice recommendations

	And Moreover, the results show that all employees can be effective leaders, particularly by using their initiative to solve problems and by teaching and supporting colleagues And The present study indicates that it is suitable for examining pandemic-imposed change in health services, particularly EMS The findings suggest this as this was the framework used and as the study was conceived around this framework, what knowledge would one expect to come out of this work other than to confirm such a preconceived position? The data presented does however present a different story, and I would encourage the researchers again to consider if these were forced rather than emergent findings? The researchers do however highlight such a position in their limitations by stating: However, a potential limitation of this study was that it used deductive thematic analysis, which risks overlooking themes that do not match the selected framework. In their mitigations they tell how the entire dataset was coded, including responses that did not seem relevant to the framework which led to the identification of a seventh QI driver, 'Infrastructure for quality'. I do not consider this sufficient mitigation to offset the fundamental issues with this approach. They declare another potential limitation highlighted above, where that the participants could have felt uncomfortable providing negative feedback, despite being assured that the interviews were confidential and that the interviewers were external to NAS. Again their mitigation does not provide assurance on this. This study may indeed provided novel insights on staff experiences in a national ambulance service in the midst of a pandemic. The researchers highlighted calls for research that improves the rigor and quality of improvement practices and evaluations. There were many strengths within the methods and methodologies and this study had an opportunity to present a rigorous exploration of this subject. I do however feel that the study is fundamentally undermined by adopting the HSE Framework for Improving Quality as the findings seem forced rather than emergent. I would encourage the researchers to explore the data closer without the constraints of this framework and allow themes to emerge. 1. Fitzsimons J. Quality and safety in the time of Coronavirus: design better, learn faster. Int J Qual Health Care. 2021;33 2. Lipworth W, Gentgall M, Kerridge I, Stewart C. Science at Warp Speed: Medical Research, Publication, and Translation During the COVID-19 Pandemic. J Bioeth Inq. 2020 Dec;17(4):555-561. doi: 10.1007/s11673-020-10013-y. Epub 2020 Aug 25. PMID: 32840838; PMCID: PMC7445735. 3. Johnson RB, Onwuegbuzie AJ. Mixed methods research: a research paradigm whose time has come. Educ Res. 2004;33:14–26.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1 - Dr. Yusuke Katayama, Osaka University

Comments to the Author: Thank you for the opportunity of reviewing this manuscript about Irish National Ambulance Service response to the COVID-19 pandemic. I was very interesting in reviewing this manuscript, but this manuscript had many problems for publication.

Thank you for your comments. We have made the changes you suggested to the manuscript.

Major point - First of all, most of the readers do not fully understand the COVID-19 pandemic in Ireland. Therefore, the authors should describe the COVID-19 pandemic in Ireland in detail.

As recommended, we have added a description of the COVID-19 pandemic in Ireland to the Introduction section (Lines 84-97).

Secondly, the authors should describe how many COVID-19 response teams there are in Ireland. Without this data, it is impossible to determine if the number interviewed in this study is appropriate. Before conducting this study, did you make an estimate of the minimum number of people who the author would need to interview? Even if the number of COVID-19 response team members is small, the number of people interviewed was small, and the authors should probably have more participants in this study in order to generalize the results.

Thank you for this comment. We have now clarified in the paper that there was just one COVID- 19 Response Room that was responsible for managing all requests for COVID-19 testing and that was provided by the National Ambulance Service of Ireland (Lines 102-106). We have noted that we recruited approximately half of this team for this study (Lines 197-198). In addition, we have provided additional details about our sampling strategy (Lines 176-182), which aimed to meet the qualitative criteria of saturation and maximum variation, rather than the quantitative criteria of generalisability.

Third, although it was stated that the COVID-19 response team includes a general practitioner, there was no general practitioner among those interviewed in this study. All professions should be included in the study in order to evaluate the feedback from the interviews.

We have amended the paper to acknowledge that a limitation of this study was that a general practitioner was not interviewed (Lines 565-568). We have stated that we attempted to recruit a medical doctor but they were unavailable due to work commitments during the third wave of the pandemic in Ireland (Lines 198-200).

Finally, the authors did not clearly describe the conclusion. If the conclusion is not clearly described, the readers cannot understand what the authors would like to convey in this study. The authors should clearly describe what they would like to convey, including what is important in such a health crisis.

As recommended, we have added a conclusion section to the paper that describes the messages we wish to convey (Line 621).

Minor point - In Page 8, the phrase "it has been reported in line with the Standards for Reporting Qualitative Research" should be written at the end of "design" section.

We have made this change (Line 140).

In Page 8, the authors described "The full protocol has been published previously.", but the authors should cite the literature about the full protocol.

As requested, we have cited the published protocol (Lines 123-125).

Reviewer: 2 - Mr. Nigel Rees, Welsh Ambulance Service NHS Trust

Comments to the Author: Congratulations to the team in gathering evidence during this time and developing the body of knowledge in this area. There is a good discussion around the role of QI and improvement science methodologies. It is also argued that the COVID-19 has prompted the development of quality improvement (QI) methods and tools suited to evaluating pandemic-imposed change, such as rapid learning cycles and after-action reviews. Despite this it is difficult to find evidence of this within the papers cited and those which are cited such as Fitzsimons [1] appear to lack such scientific rigor and include opinion pieces. It is also difficult to determine the scientific quality of such approaches when the researchers highlight how in times of crisis, QI approaches must be adapted and streamlined so that improvements are implemented quickly without unnecessarily burdening strained services and staff. The paper may benefit from considering the validity of such approaches and implications of adapting or streamlining such approaches? A counter argument to such a position has been made in many articles which report poorly governed projects and low-quality methodologies employed during this time such as discussed by Lipworth et al (2020) [2].

Thank you for your valuable feedback. We have taken time to reflect on your comments and have made various changes to the paper as a result.

We agree that the paper would benefit from considering the validity and quality of QI approaches

that have been adapted and streamlined for a crisis. Therefore, we have added this point to the Introduction (Lines 81-83), Discussion (Lines 536-546), and Conclusion (Lines 633- 636). We have cited several additional papers, including Lipworth et al. (2020).

The researchers tell how the research paradigm was pragmatism, and they report how the study was a component of a mixed methods evaluation of the NAS response to the pandemic. Whilst they provide Johnson et al (2004) [3] as a reference for this, it would help to be cited on the full study. Was the protocol prospectively published? Or are any outputs in the public domain?

As recommended, we have now provided the citation for the published protocol of the mixed methods evaluation (Lines 123-125) and clarified that it was first published prior to this qualitative study being carried out. There are no other outputs in the public domain at present.

There is a good description on the research methods, development, and review of the interview schedule, and aim for data saturation. There is also a good description of the sample. I was however concerned with participants being recruited via a study invitation email sent by a NAS manager to former and current staff. Has the team considered their anonymity in this process and if participants were indeed open to discuss freely?

Thank you for this feedback. We have now included this point in the limitations section (Lines 568-584). We have also included additional information about this aspect of the recruitment process in the Methods section (Lines 183-189). We took a number of steps to ensure that participants would remain anonymous and would feel comfortable speaking openly during the interviews but we have acknowledged that they may have been reluctant to do so nevertheless.

There is a good description of background of researchers, which is important in their chosen methodology. They have also told how they maintained field notes and reflexive notes throughout. The use of Braun and Clark is also appropriate. This along with the peer debriefing enhanced the trustworthiness of the analysis. The researchers however mentioned how the interview schedule was reviewed and, where necessary, refined following each interview (Supplementary File 2). It would help if they could justify and describe further methodology used.

As recommended, we have added a description and justification for the refinement of the interview schedule to the Design section (Lines 126-136) and Procedure section (Lines 214-218).

The researchers told how the research was underpinned by the HSE Framework for Improving Quality and that it was used to examine the utility of this framework in a pandemic

setting and inform the preparation of health services, particularly Emergency Medical Services (EMS), for future crises. Whist it seems acceptable to draw on this framework, the study in my view is significantly undermined by the influence of this framework. I would encourage the researchers to consider and provide a discussion on the prominent role of this framework and implications of emergence of themes versus forcing the data and confirming knowledge within this framework. For instance, many of the quotes provided do not seem to fit this framework but are coded within these themes.

Thank you for this feedback. We have made several amendments to the manuscript in light of your comments. Firstly, in the Methods section (Lines 222-240), we have provided a more detailed explanation of the specific qualitative analysis technique that we used: Braun and Clarke's (2006; 2019; 2020) reflexive thematic analysis procedure, which is widely used in qualitative research. We have reported that this procedure is carried out at a point on the continuum between a primarily inductive analysis and a primarily deductive analysis. In our study, we used a primarily deductive approach given the prominent role of the Framework for Improving Quality. Deductive analysis is used to evaluate, explicate, amend, or challenge an existing framework, rather than to simply endorse it or to force the data to fit within it. In addition, the importance of the framework can increase or decrease as the analysis progresses depending on the researcher's assessment of its relevance and utility.

Secondly, in the Methods section (Lines 169-172 and Lines 241-263), we have provided additional information about the specific role of the Framework of Improvement Quality in the analysis. We clarified that the aim was to explore experiences and perceptions of the quality improvement drivers from the framework in the context of a pandemic-response room. The framework could be amended as part of the analysis to ensure that the dataset was adequately captured, especially as it had not been previously applied in this context. For example, new drivers could be added or existing drivers could be re-contextualised or removed.

Thirdly, in the Discussion section, we have now provided a more detailed discussion of the implications of using a deductive approach, as opposed to an inductive approach (Lines 585- 620). We acknowledge that the deductive approach has limitations. In particular, we acknowledge that there is a risk that important concepts or themes that do not match the chosen framework will be overlooked, even when steps are taken to mitigate this risk. We also note that deductive analysis has potential strengths and is an established and accepted approach in qualitative healthcare research. In contrast with alternative qualitative approaches, themes are not regarded as having emerged from the data in the Braun and Clarke procedure. Instead, the researcher is viewed as having an active role in constructing the themes and as being influenced by their preconceptions and knowledge of the relevant literature and theory, even when an inductive approach is used.

Finally, we have made several amendments to the Results section based on your comments (Line 264). We acknowledge that we needed to provide a more convincing account of certain themes,

including providing appropriate supportive evidence/quotes for those themes. We also needed to clarify that the themes represent the participants' interpretations, perspectives, and experiences of the QI drivers from the framework in the context of a pandemic response room (Lines 171- 172 and 260-263). Therefore, the content of the themes differed somewhat from the original content of the framework. The specific changes made to the Results section are outlined below.

The Leadership for Quality theme is one such example as the following quote talks about management, performance, and wellbeing, yet there is little consideration of these issues within the discussion or how they relate to leadership?

"How it's managed: That's the crux of it all…You can have amazing workers…but if it's not managed with respect and…caring, people won't…get the job done…because…it's just like a pressure cooker," (Interviewee 7).

The 'Leadership for Quality' section has been revised in light of your comments (Lines 268-307). The above quote has been replaced with a more relevant quote. The quotes presented are now discussed more clearly in terms of how they relate to leadership.

The following quote was also under the theme of Use of improvement methods. Despite this, there is little evidence in the quote of improvement methods/science/models or how they were adapted, but rather it seems to refer to leadership:

Many reported that the response room manager provided strong leadership:

Nevertheless, improvements were possible: "The clinical team…towards the end of the first peak...created a few standard operating procedures and we created lists…with all the

important contacts [and] what we thought worked, what we thought didn't work...That was a really invaluable document…if they were training anyone," (Interviewee 8).

The following quote also undermines some of argument made in the paper around improvement and rigorous QI approaches:

"During the first wave, there was very little time to…look back and see how things could be

improved. It was more so a case of…just fighting against the wave…certainly from our level," (Interviewee 9). **N**

Many authors have highlighted how much of the experiences of the COVID-19 pandemic had been endured by others during previous pandemics, were somewhat predictable and within the published literature. Rather than having little time to look back, it has been argued that reflecting on/reviewing literature from previous pandemics should have been essential.

We reflected on your feedback and have subsequently amended both the Results section and Discussion section.

We agree that the data show that there was little use or adaption of rigorous improvement methods/science/models in the COVID Response Room. The improvements made in the room during the study period were largely informal and ad hoc. We have revised the Discussion section (Lines 531-550), Conclusion section (Lines 629-636), and the 'Use of Improvement Methods' section (Lines 369-399) to clarify this. In addition, the phrase "*Many reported that the response room manager provided strong leadership*" is not included in the latter section.

We agree that reviewing literature from previous pandemics is crucial. Therefore, we have added this point to the Discussion section (Lines 542-544).

There is again a mismatch between the theme The measurement for quality. There is no evidence of improvement science/QI models or their adaptations within the quote. This theme and resulting discussion may therefore have been better served by exploring the lack of rigor, validity, definition, and deployment of QI models in the EMS during the pandemic?

We have amended the 'Measurement for quality' theme in light of your comments. Specifically, we have amended the Results section (Line 400-444), Discussion section (Lines 531-550), and Conclusion section (Lines 629-636) to clarify that the measurement within the response room lacked rigour and that the use or adaptation of improvement science methods and models would have been possible and beneficial.

In my view, the theme Infrastructure for quality is an example to the rest of the study themes, as they appear to suggest that it emerged from the data rather than being forced into a framework which already existed. Again however, much of the quotes may relate to something else other than infrastructure, such as communication and IT?

We have changed the name of this theme to 'Information and communications technology for quality' based on your feedback (Line 473).

Within the paper many inferences do not seem to be justified by this study, such as: *In particular, the findings suggest that there are seven QI drivers that should be addressed when preparing for and responding to pandemics and similar emergencies.*

And: *The results underscore the importance of empathic and respectful leadership that* comprises regular and clear communication and consultation with staff, which aligns with best practice recommendations

And: *Moreover, the results show that all employees can be effective leaders, particularly by using their initiative to solve problems and by teaching and supporting colleagues*

And: *The present study indicates that it is suitable for examining pandemic-imposed change in health services, particularly EMS*

These statements have been amended or removed from the Abstract (Lines 33-50), Discussion section (Lines 497-512), and Conclusion section (Lines 621-638).

The findings suggest this as this was the framework used and as the study was conceived around this framework, what knowledge would one expect to come out of this work other than to confirm such a preconceived position? The data presented does however present a different story, and I would encourage the researchers again to consider if these were forced rather than emergent findings? The researchers do however highlight such a position in their limitations by stating:

However, a potential limitation of this study was that it used deductive thematic analysis, which risks overlooking themes that do not match the selected framework.

In their mitigations they tell how the entire dataset was coded, including responses that did not seem relevant to the framework which led to the identification of a seventh QI driver, 'Infrastructure for quality'. I do not consider this sufficient mitigation to offset the fundamental issues with this approach.

Thank you for this feedback. As outlined earlier, we have made major amendments to the Methods, Results, and Discussion sections to address this. In particular, we expanded the discussion of the limitations of deductive thematic analysis in the Discussion section (Lines 585- 620). We acknowledge that a key limitation (i.e. that important concepts or themes could have been overlooked) remains, despite our efforts to mitigate it. We also acknowledge that an alternative method (e.g. inductive thematic analysis, grounded theory) would likely have produced different findings that were more data-driven. However, we also state that deductive thematic analysis is a valid and established approach in qualitative research. Its aim is not to uncritically endorse a framework or to force data to fit within it but is instead to evaluate, explicate, amend, or challenge that framework. In this study, we used a deductive approach that permitted the amendment of the framework, which included adding, removing, or re- contextualising drivers, to adequately capture the experiences of healthcare workers in a pandemic setting. We have amended the Methods section (Lines 222-263) to clarify this point and to provide a more in-depth explanation of the

specific approach we used: primarily deductive, reflexive thematic analysis using the Braun and Clarke procedure.

In the Braun and Clarke/reflexive thematic analysis approach, it is argued that multiple analyses or interpretations of any dataset are possible. Therefore, what is important is that the researcher is transparent about the particular approach that they have used, including its implications and limitations, so that the findings can be interpreted and critiqued appropriately. Furthermore, in this particular approach, whether inductive or deductive, themes are not considered to have emerged from the data. Instead, the researcher is viewed as having an active role in the construction of themes at the intersection of the data and the researcher's aims, subjectivity, theoretical understanding, training, and experience.

They declare another potential limitation highlighted above, where that the participants could have felt uncomfortable providing negative feedback, despite being assured that the interviews were confidential and that the interviewers were external to NAS. Again their mitigation does not provide assurance on this.

We have expanded the Methods section (Lines 183-189) and limitations section (Lines 568-584) to address this point. We acknowledge that a limitation of the study is that, despite the various steps taken to reassure participants, at least some may still have felt uncomfortable speaking freely and providing negative feedback during the interviews.

This study may indeed provided novel insights on staff experiences in a national ambulance service in the midst of a pandemic. The researchers highlighted calls for research that improves the rigor and quality of improvement practices and evaluations. There were many strengths within the methods and methodologies and this study had an opportunity to present a rigorous exploration of this subject. I do however feel that the study is fundamentally undermined by adopting the HSE Framework for Improving Quality as the findings seem forced rather than emergent. I would encourage the researchers to explore the data closer without the constraints of this framework and allow themes to emerge.

Thank you for these insightful comments. We have made several changes throughout the manuscript to address your concerns. This includes clarifying that the themes represent the COVID Response Room employees' experiences and perspectives of quality improvement drivers from the framework. Therefore, the content of the themes may differ somewhat from the content of the drivers from the original framework. In addition, we provided a more detailed discussion of the strengths and limitations of deductive and inductive approaches to data analysis. We acknowledged that the deductive approach is theory- and researcher-driven, rather than data-driven. We also acknowledged that the limitations of deductive analysis cannot be completely mitigated, even through the use of strategies such as coding inductively and permitting amendments to the selected framework. Nevertheless, deductive analysis is an established and valid approach, provided that researchers are transparent about its use and its shortcomings.

VERSION 2 – REVIEW

REVIEWER	Nigel Rees Welsh Ambulance Service NHS Trust, Pre Hospital Emergency Research Unit
REVIEW RETURNED	13-Dec-2021

GENERAL COMMENTS	Thank you once again for giving me the opportunity to review this paper. This is a very interesting paper and I again congratulate the team for conducting this study, preparing the paper and responding to reviewers comments within the challenging context of the pandemic. The authors have provided detailed clarification of the peer reviewer comments which are also reflected in the paper. The paper will make a valuable contribution to the body of knowledge. Regards Nigel